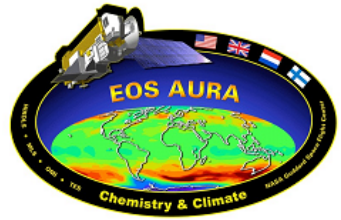
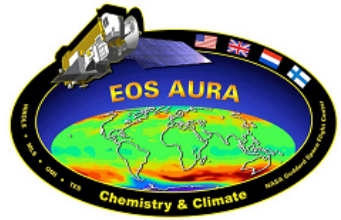


Aura Science Meeting
Data Systems Working Group
HIRDLS SIPS Status
October 1, 2007

Vince Dean, Brendan Torpy, Greg Young
Univ. of Colorado, Boulder

Cheryl Craig
NCAR

Overview

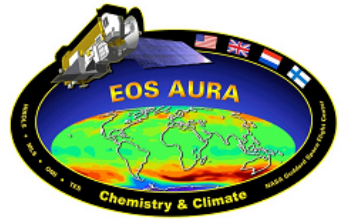


- **Status of processing—science and data users**
  - **Software versions**
  - **Delivery to AVDC and to DISC**
  - **Next steps**
- **SIPS system—developers and data managers**
  - **Lessons learned**
  - **Support for multiple processor versions and experimental runs**
  - **File and processor versioning**

Since September 2006

- Delivered v2.02.07 HIRDLS Level 2 product to Goddard DISC (314 days of interest)—our first public release.
- Installed 106 versions of science processors in HIRDLS SIPS for experimental and production processing.
- Processed 897 days of data (full mission) with v2.04.08 software.
- Processing full mission with v2.04.09 software; 679 days completed so far.
- Supported ESDIS evolution and mini-MOSS tests:
  - S4PA mini-MOSS – December 2006
  - TADS mini-MOSS – April 2007
  - GEOS-5 v1.0 mini-MOSS – September 2007
- Hardware acquisitions:
  - Installed 32-processor SGI Itanium system
  - Replaced SIPS PC cluster with new rack-mount systems.

Obstruction

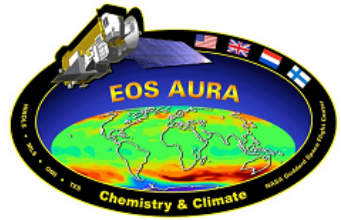


- An obstruction partially obscures the light path within the instrument.
- Algorithms continue to be developed to correct for the effects of the obstruction.
- This highly experimental software development has impact on:
  - Plans for processing and data release.
  - SIPS operations and software systems.

Recent Data Releases

- **v2.04.08**
  - Processed 350 selected days of interest.
  - Basis for several publications.
  - Bug found during QA of v2.04.08.
  - Review shows no impact on conclusions in publications, but warranted reprocessing.
- **v2.04.09**
  - Fixes the bug.
  - Basis for other publications.
  - Plan to process entire mission: 679 days so far.
  - Plan to release to AVDC and DISC for public distribution in October.
- **DISC now supports arbitrary strings for version numbers, like**
  - **“2.04.09”**

Next Release



- **Driven primarily by advances in algorithm development.**
- **Will incorporate:**
  - **Refinements to correction for obstruction**
  - **Possible scientific enhancements**
  - **GEOS-5 1.0**

GEOS-5

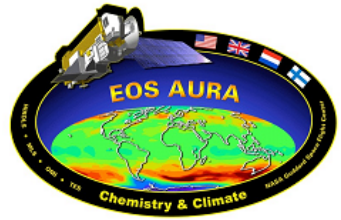
- Our recent processor versions use GEOS-5 v0.1 for temperature line-of-sight gradients:
  - 2.04.08
  - 2.04.09
- GEOS-5 1.0 can have a significant impact on HIRDLS data product; this new release will necessitate a new version of HIRDLS products.
  - Will wrap reprocessing using GEOS-5.1 along with further data processing refinements
- When GEOS-5 v0.1 stops production, we will discontinue daily processing with v2.04.09 until our next major software release.

ESDIS Evolution

- **Significant, but bearable, impact on SIPS operations.**

- **HIRDLS appreciates quick response of Goddard DISC as S4PA and our needs evolve.**
  - **Enhancements to Machine Request Interface will provide valuable tool for data reconciliation and routine operations.**

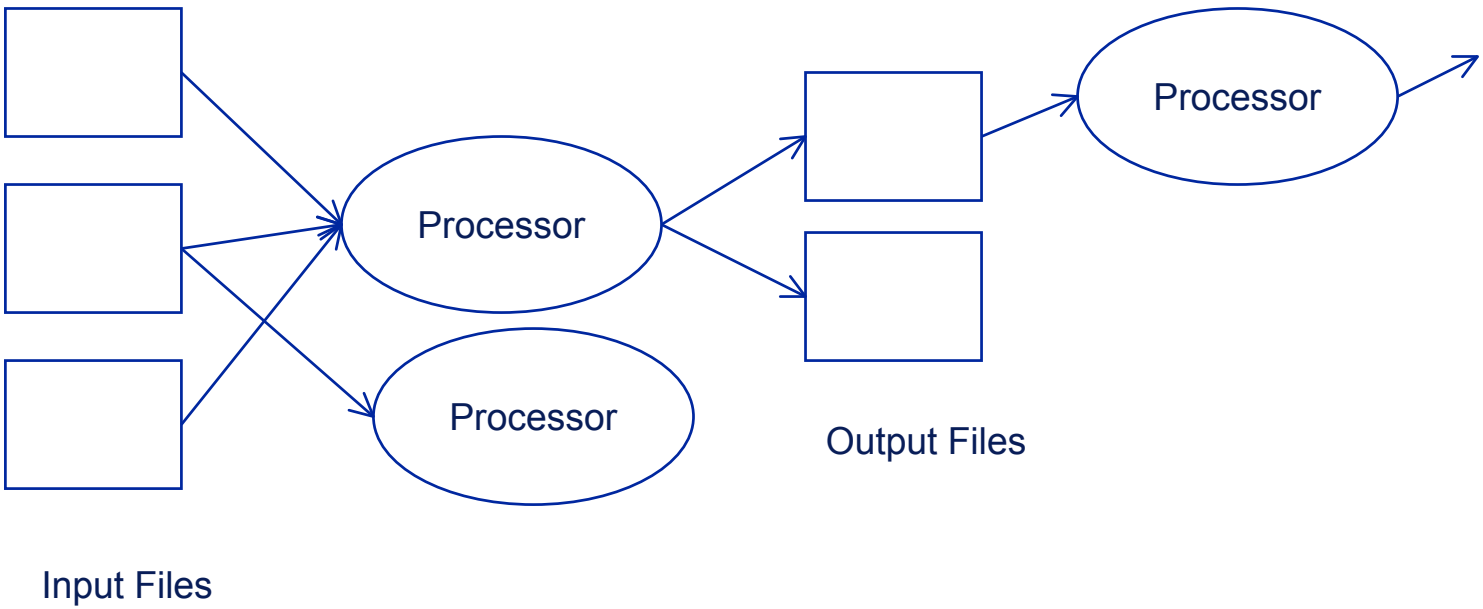
SIPS System—Software and Operations



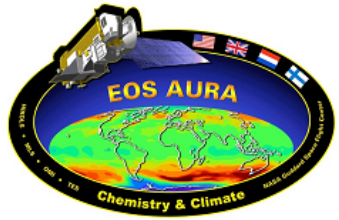
- **Embrace change—embrace experimentation.**
- **Supported by some fortunate design decisions.**
- **Some examples:**
  - **Issues—conflicting forces**
  - **Resolution**
  - **Consequences**

Processing History

- Processing history is a complex graph of data files and processing steps.
- Metadata is preserved from every run.
- Every product is traceable back to raw inputs.
- Supports experimentation, but complex to manage.

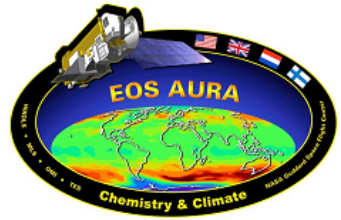


Experiments in a Production System



- **Forces:**
  - Many experimental processor versions
  - Manual bookkeeping strained by complexity
- **Resolution:**
  - Embrace the complexity
  - Accept the many experimental processors into the production system.
  - Run experiments in production system
- **Consequences:**
  - Reliable audit trail.
  - Efficient bulk re-processing.
  - Version tracking is more complex.
  - More complex work for SIPS operator.

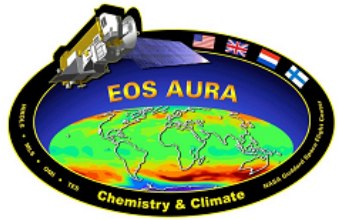
Specifying Processing Tasks



- **Forces:**
  - **We frequently run tests where only a few processors have changed.**
  - **We don't want to rerun previous steps.**
  - **It is inconvenient to customize the processing specification according to the state of processing on each day.**
- **Resolution:**
  - **Operator specifies desired outcome—required steps.**
  - **System identifies which steps have already been run and does not repeat them.**
  - **Like a Unix "make file".**

Make it so!

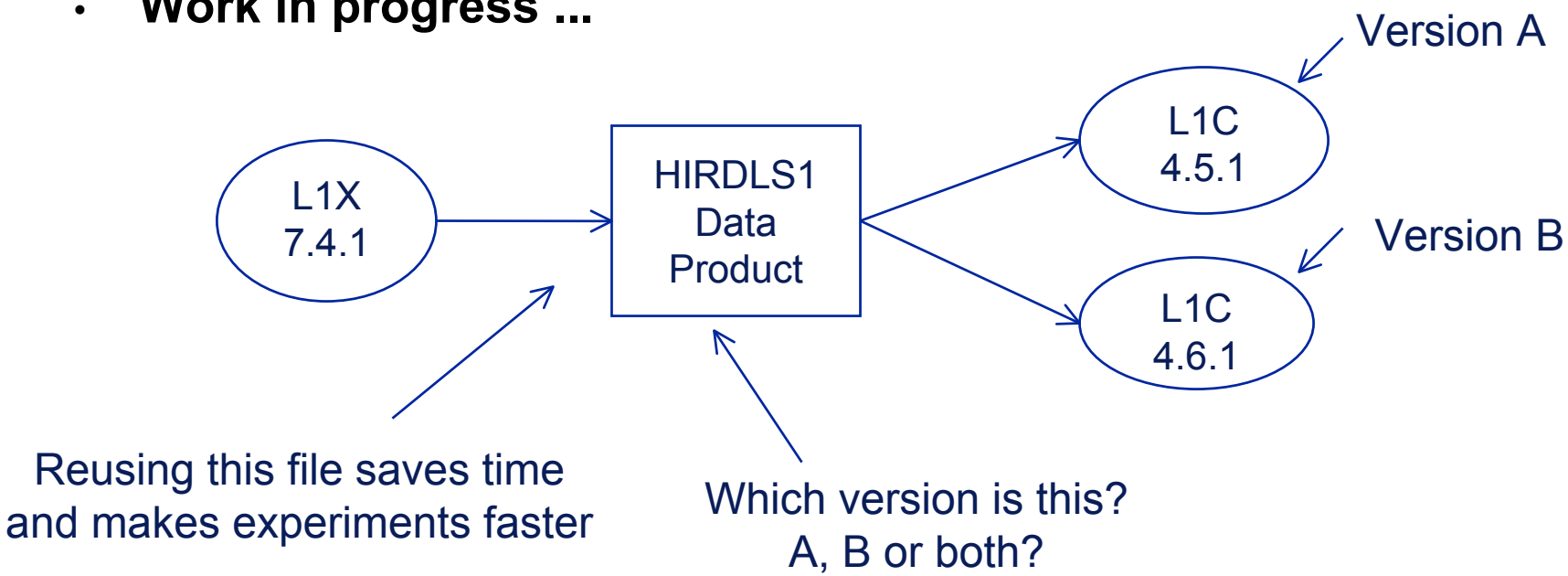
Sharing Experimental Versions



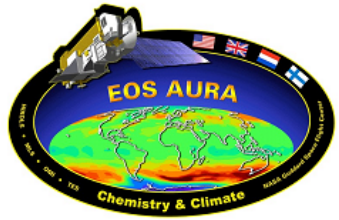
- **Forces:**
  - HIRDLS SIPS experimental processing history is complex
  - Users, especially external users, need easy access to the right data
  - Risk of using wrong data files
  - Prefer a simple, hierarchical file structure
- **Resolution:**
  - HIRDLS “data pool”
  - Web server with selected files in a simple hierarchy.
  - External program periodically copies selected data products to the web server.
  - Creates a directory for each data version
  - External users download new products automatically with `wget`.

Identifying File Versions

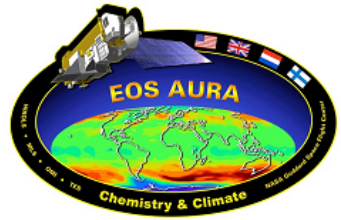
- **Forces:**
  - **We release one component at a time.**
  - **Want to have unique, informative labels for all products.**
  - **Want to reuse old processing products in new sequences.**
- **Work in progress ...**



Monitoring Many Systems



- **Forces**
  - System has many components
  - Failures may not be immediately obvious
  - Each has a different mechanism for reporting problems
  - Email notifications can be overwhelming
- **Resolution**
  - External monitoring tool
  - Nagios: www.nagios.org
  - Email notification and web interface
- **Consequences**
  - State of system is readily visible
  - We are aware of problems before the users are
  - Some outages cause multiple notifications



Apache Tomcat/5.5.20 - Error report

Current Network Status
Last Updated: Wed Sep 26 12:16:43 MDT 2007
Updated every 90 seconds
Nagios® - www.nagios.org
Logged in as *vdean*

View History For all hosts
View Notifications For All Hosts
View Host Status Detail For All Hosts

Host Status Totals

Up	Down	Unreachable	Pending
31	1	0	0

All Problems	All Types
1	32

Service Status Tota

Ok	Warning	Unknown	Critic
67	3	0	6

All Problems	All Ty
9	76

Display Filters:
Host Status Types: All
Host Properties: Any
Service Status Types: All Problems
Service Properties: Any

Service Status Details For All Hosts

Host	Service	Status	Last Check	Duration	Attempt	Status Information
hcl	PING	CRITICAL	09-26-2007 12:13:09	0d 9h 11m 27s	1/4	CRITICAL - Host Unreachable (hcl.acd.ucar.edu)
hir1	disk_slash_sandbox	CRITICAL	09-26-2007 12:13:14	0d 9h 11m 22s	4/4	CHECK_NRPE: Socket timeout after 10 seconds.
	disk_slash_sips	CRITICAL	09-26-2007 12:15:09	1d 16h 56m 52s	4/4	CHECK_NRPE: Socket timeout after 10 seconds.
hirdlssips	Datapool Browse	WARNING	09-26-2007 12:11:23	0d 0h 15m 20s	4/4	HTTP WARNING: HTTP/1.1 404 /sips/datapool/
	Datapool README	WARNING	09-26-2007 12:13:18	0d 0h 18m 25s	4/4	HTTP WARNING: HTTP/1.1 404 /sips/datapool/README.html
	HIR0SCI Browse	WARNING	09-26-2007 12:11:28	0d 0h 15m 15s	4/4	HTTP WARNING: HTTP/1.1 404 /sips/public/HIR0SCI/
hirz1	Trend log file	CRITICAL	09-26-2007 12:13:33	0d 6h 30m 18s	4/4	FILE_AGE CRITICAL: /hirtrend/trend.log is 37707 seconds old and
lvl	Trend plot	CRITICAL	09-26-2007 12:12:02	0d 6h 29m 34s	4/4	CRITICAL - Last modified 10:27:21 ago
mail-acd	ACD Mailbox	CRITICAL	09-26-2007 12:12:06	0d 3h 34m 30s	4/4	gdaac_mailbox CRITICAL: 41 messages in inbox

9 Matching Service Entries Displayed